REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

AD-A238 507

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY	2. REPORT DATE 90	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE

Combining CL (confidence limits) and PL (partial limits) displays

5. FUNDING NUMBERS

DAAL03-88-K-0045

6. AUTHOR(S)

John W. Tukey

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 26098.21-MA

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

There are a variety of circumstances where we might want to show two (or even three) intervals of uncertainity around each measured data point. The most frequent involves a wider interval (between SCL's) for comparing values with a numerical scale and a narrower interval (between SPL's) for comparing results with one another. In that circumstance, the use of double open triangular arrowheads (a "DOTAR" display) is particularly useful.

Having chosen a graphical style, we need quantitative insight and numerical tables to implement our choice. In particular, we need to recognize the distinction between "severe simultaneous" and "tight simultaneous" and the possibility of finding, and using, severe simultaneous SCL's and SPL's at large exception rates. Numerical tables that outline quantitative behavior are provided.

14. SUBJECT TERMS

Apertures, Bonferroni, Confidence limits, Dotars, Double open triangular arrowheads, Graphical display, Partial limits, Severe limits, Simultaneous limits, Studentized range, Tight limits

15. NUMBER OF PAGES
27

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Combining CL (confidence limits) and PL (partial limits) displays

John W. Tukey
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

Technical Report No. 300
Department of Statistics
Princeton University
Princeton, NJ 08544-1000

June 1990

91-05553

TABLE OF CONTENTS

LIST OF EXHIBITS

Combining CL (confidence limits) and PL (partial limits) displays

John W. Tukey

Technical Report No. 300

Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

ABSTRACT

There are a variety of circumstances where we might want to show two (or even three) intervals of uncertainity around each measured data point. The most frequent involves a wider interval (between SCL's) for comparing values with a numerical scale and a narrower interval (between SPL's) for comparing results with one another. In that circumstance, the use of double open triangular arrowheads (a "DOTAR" display) is particularly useful.

Having chosen a graphical style, we need quantitative insight and numerical tables to implement our choice. In particular, we need to recognize the distinction between "severe simultaneous" and "tight simultaneous" and the possibility of finding, and using, severe simultaneous SCL's and SPL's at large exception rates. Numerical tables that outline quantitative behavior are provided.

KEYWORDS:

Apertures, Bonferroni, Confidence limits, Dotars, double open triangular arrowheads, Graphical display, Partial limits, Severe limits, Simultaneous limits, Studentized range, Tight limits.

June 11, 1990

Combining CL (confidence limits) and PL (partial limits) displays

John W. Tukey

Technical Report No. 300

Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

1. Introduction.

When we have measurements - - apparent values and estimated uncertainties - - on a number of candidates (treatments, locations, circumstances), two sorts of expression of uncertainty are frequently useful or needed:

1) confidence limits (CL's), which express the range of uncertainty for the "true value" for each *individual candidate*, and

2) partial limits (PL's), which are stated for individual candidates, but which are combined in *pairs* to express uncertainty about *differences* between pairs of candidates.

As we shall see, there are a variety of ways to express the chosen degree of uncertainty - - the degree chosen to convert flexibly-usable information into specific limits. Whatever rule may be chosen, the following point can be expected to apply (at least for small and moderate uncertainties):

Given ranges of "reasonable possibilities" for two specific candidates (with uncorrelated errors), while it is individually reasonable, say, that the "true value" for candidate A is close to the upper end of its reasonable range, and that the "true value" for candidate B is close to the lower end of its reasonable range, it is *not at all equally reasonable* that *both* of these things have happened together. As a consequence, we get unduly wide ranges for *differences* if we add up the half-ranges of uncertainty for the individual candidates.

2. The basic double-open-triangle-arrow (dotar).

It is important to the practical solution of this problem that it is the shorter intervals that are to be combined in pairs - - with special attention to overlap or non-overlap. This calls for making the ends of the shorter intervals (PL's) more easily comparable with each other. (The ends of the longer intervals need only to be compared with the scale.)

Exhibit 1 shows how a double-headed arrow, whose arrowheads are open triangles, can be used to display both confidence limits (for individual candidates) and partial limits (to be combined for comparing pairs of candidates). The horizontal bases of the triangles provide the firm reference that is needed to compare pairs of individual candidates. The tips of the triangles are specific enough to give a good picture of overall ranges of reasonable values.

---

exhibit 1

about here

Until a better procedure is invented/discovered, this double-triangle-headed-arrows form of display seems to be the display of choice.

3. Individual or simultaneous limits.

The need for both CL's and PL's is greatest where there are several - - and even many - - individual candidates. As a consequence, problems of multiplicity have to be faced.

If, for instance, we have 50 candidates, and (for simplicity) the observed values all deserve the same standard errors, (which has been determined on the basis of many observations - - many degrees of freedom), and if we set confidence limits 2s above and 2s below the observed values, the chance that any one - - externally identified - - candidate's "true value" will fall outside its confidence interval (the range between the confidence limits) is close to 5%. The average number outside will be 50 times 5%, namely 2.5 (= 250%). If we have many different users, each of whom needs to ask about only one particular candidate, our attention is almost sure to go to the 5% per individual - - rather than to the "250% per study". But if users habitually

exhibit 1

How to express both CL's (confidence limits) and PL's (partial limits)
with double-headed arrows
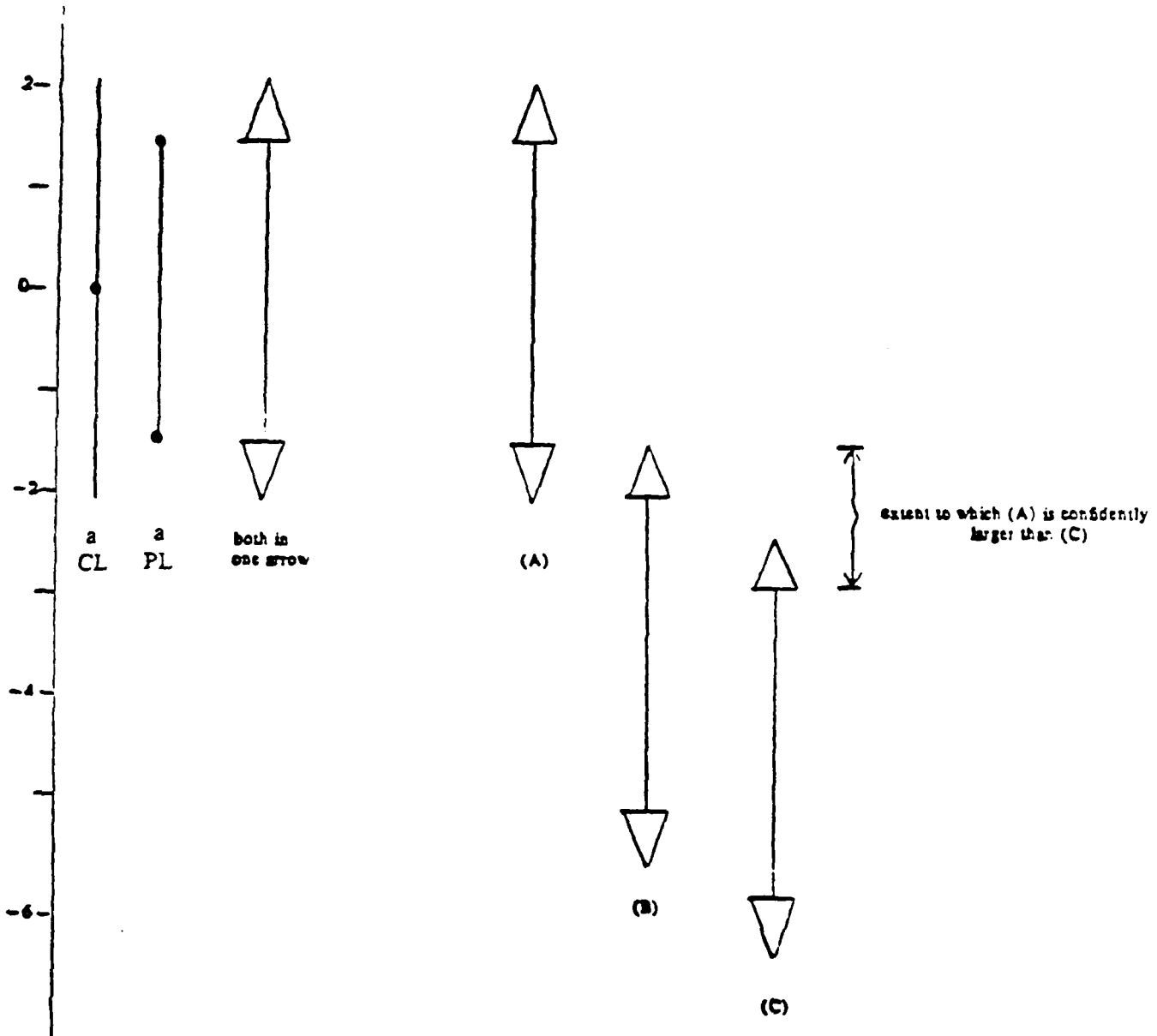
seek out the unusual, we can hardly fail to be seriously concerned about the 250% per study.

A Bonferroni approach, perhaps one in which we spend 1/50th of 5% - - 0.1% - - on each of 50 candidates, is the extreme response to such concerns, leaving us with 0.1% per individual and 5% per study. In such a case we would refer to *simultaneous confidence intervals*, since we could be confident - - with only 5% diffidence - - that *all* the individual "true values" are in the corresponding (simultaneous 5%) CL's.

The situation is more extreme for the comparisons of pairs of candidates. If we spend 5% on each pair from our 50, the average number of errors amounts to 1225(5%) = 6125% - - or an average of 61.25 misses, 61.25 pairs for which the implied CL for their difference would not include its "true value". (Because we have asked for the *average number* of errors, which is the sum of the probabilities of individual errors - - without regard to dependence or independence - - statistical dependence between certain pairs makes no difference.)

Again if each user only wanted to think about a single pair, chosen without regard to the observed values, maybe "5% individual" would dominate our thinking, leaving the "6125% simultaneous" not too serious. But has any of us ever seen such a user - - even one such? I doubt it. When we come to pairwise comparisons, most users will shop around looking for interesting comparisons. They need much stronger multiplicity control, and a Bonferroni adjustment, even if it involves $5\%/1225 \doteq .004\%$ individually - - to get 5% simultaneous, we must do something of this magnitude - - does not seem out of place.

If we "just love" 5%, then, there are at least three sets of limits we may want to look at for each of k candidates (* is a multiplication sign):

1) 5% simultaneous PL's (SPL's), which we can find as
$$\pm [5\%/(k(k-1)/2 \text{ point of } |t| / \sqrt{2} = 5\%/k(k-1) \text{ point of } t/\sqrt{2}] * s$$

2) 5% simultaneous CL's (SCL's), which we can find as
$$\pm (\text{the two-sided } 5\%/k \text{ point of } t) * s = \pm (\text{the one-sided } 2.5\%/k \text{ point of } t) * s$$

3) 5% individual CL's (ICL's), which we can find as
$$\pm(\text{the two-sided } 5\% \text{ point of } t) * s = \pm (\text{the one-sided } 2.5\% \text{ point of } t) * s$$

where s is the appropriate standard error. (Details are discussed in an appended section.)

We have already dealt with displaying PL's and CL's together - - we can and should use this technique for the simultaneous limits (SPL's and SCL's). The remaining pair of limits (ICL's) can probably be displayed in terms of a thicker (or doubled) bar in the center of the arrow, as illustrated in exhibit 2. If we really need *three* sets of

June 11, 1990

limits in a single display, this is the best approach to date, one that can probably be used without undue misinterpretation.

<div align="center">

---
exhibit 2

<u>about here</u>
</div>

4. Bonferroni

The Bonferroni approach to CL's adopted above says that if we wish to constrain the average total exception rate for k inquiries to 5%, where each failure of a "true value" to fall in its CL is an *exception*, it suffices to constrain each CL to an exception rate of $(5/k)$%. Thus the simultaneous-5% CL's are at

$$SCL's = \text{point estimate} \pm |t_v|[(5/k)\%]*s_v$$

where s_v^2 is an estimate of the variance of the point estimate deserving v degrees of freedom, and $|t_v|[(5/k)\%]$ identifies the two-sided $(5/k)$% point of t on v degrees of freedom, which is equal to $t_v[(2.5/k)\%]$, the one sided $(2.5/k)$% point of t on v degrees of freedom. We will refer to the multiplier of $\pm s_v$, which is the critical value of the corresponding statistic, as the coefficient of the chosen limits.

If we ask about all pairs of candidates, we are making $k(k-1)/2$ inquiries, and the simultaneous-5% limits for differences are at

$$\text{observed difference} \pm |t_v|[(10/k(k-1))\%]*s_v\sqrt{2}$$

where v and [ ] are as before, and $(s_v\sqrt{2})^2 = 2s_v^2$ is the squared standard error of each of the $k(k-1)/2$ differences. The partial limits are *half* as far as this from the individual values - - permitting a maximum displacement of $(1/2) + (1/2) = 1$ times the $\pm$ required for the differences. Thus, since $\sqrt{2}/2 = 1/\sqrt{2}$,

$$SPL's = \text{point estimate} \pm \frac{|t_v|[(10/k(k-1))]}{\sqrt{2}}s = \pm \frac{t_v[(5/k(k-1))]}{\sqrt{2}}s$$

5. Two kinds of simultaneous

All these sets of limits control the average total (= total average) error rate. We shall label them with the word "severe". If we wish to say slightly less, we can control the chance of one or more errors, perhaps keeping this chance at 5%. We shall label this with the word "tight". *Because more than one error will sometimes occur in a*

<div align="center">

June 11, 1990
</div>

exhibit 2

Modified double-open-triangular-arrow (DOTAR) display: triangles show
simultaneous-5% PL's and CL's; doubled bars show individual-5% CL's.
(shapes appropriate for $k = 2$, 5 and 10 candidates and 10 degrees of freedom, and for
$k = 10$, 20, 50 and 100 candidates with 100 degrees of freedom)
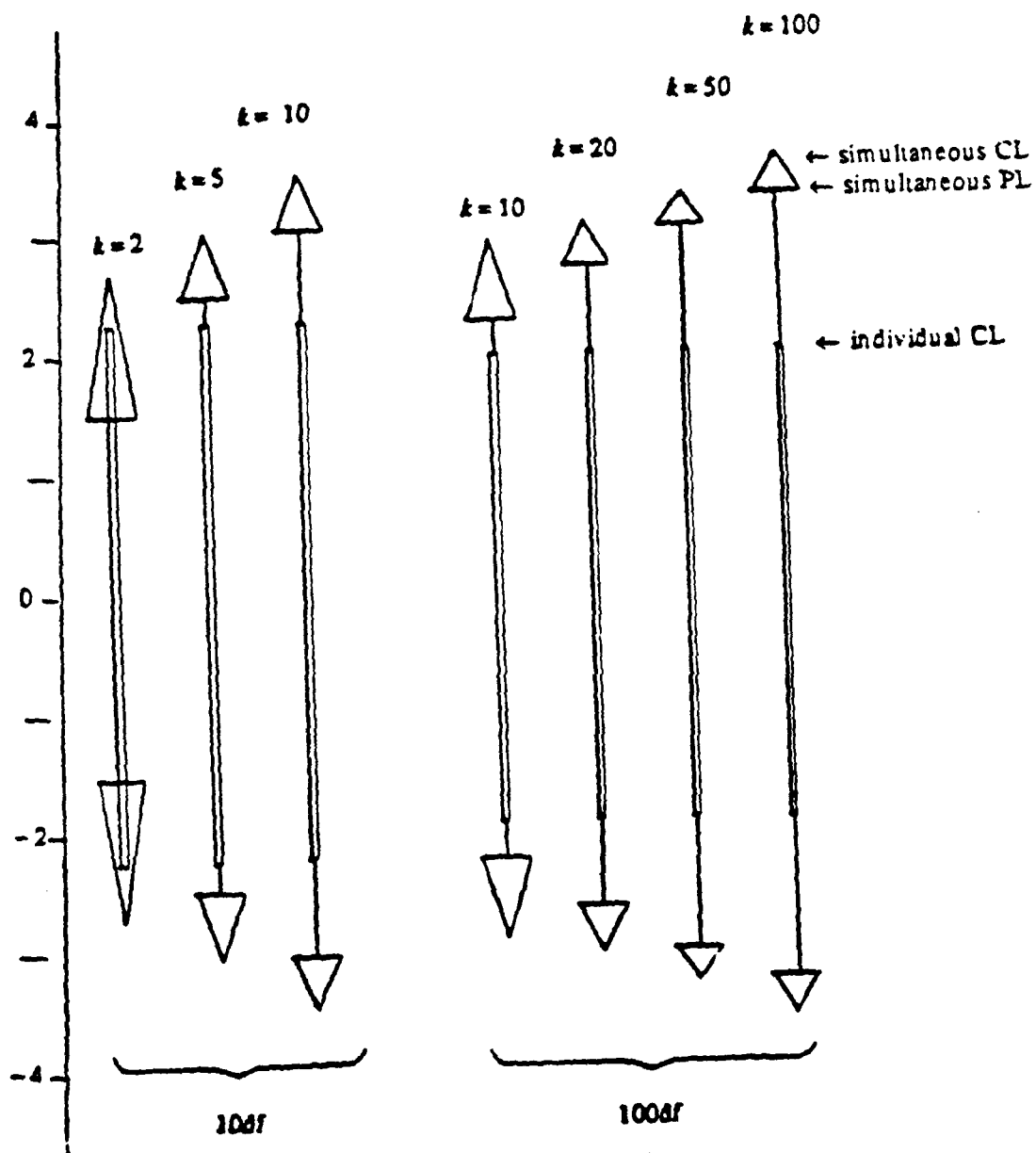
NOTE: Intrusion of the long doubled bar on the triangle only occures for $k \le 4$.

given study, the total *average* exception rate for a *tight* simultaneous procedure will be somewhat greater than 5%. Compared to the large effect of choosing between "simultaneous" and "individual", the effect of choosing between "severe" (e.g. Bonferroni) and "tight" is not too important.

Some sources refer to severe as "per batch" for SCL's and "per family" for SPL's; the same sources are likely to use "batchwise" for tight SCL's and "familywise" for tight SPL, or to call the use of the Studentized range (= 2*SPL) an "allowance".

Tight SCLs call for percentage points of the Studentized maximum modulus distribution (which is reasonably well tabulated, cp Nemenyi 1953, Pillai and Ramachandram 1954, for $k \leq 20$). Tight SPL's call for percentage points of the Studentized-range distribution (which is well tabulated, Harte. 1969), for $k \leq 400$; less completely but more widely available in Harter 1961 or Miller 1981).

So long as what has been done is made clear, for SCL's or SPL's, choice between "severe" and "tight" can probably be safely left to individual judgment.

6. Tables for 95% severe simultaneous confidence

For a mixture of gaining some convenience in use and gaining some insight into how severe SPL's and SCL's compare quantitatively, exhibits 3 to 5 give values of

$$t_v [(2.5\%/k))]$$

and

$$t_v [5\%/k(k-1)]/\sqrt{2}$$

for chosen values of v/k (for starting with 1, 1.5, 2, 3 , . . . , for $k \leq 100$, but confined to $v/k = 1$ or ∞ for $k > 100$. This form of tabulation:

- leads to rather slowly changing values - - so "eye interpolation" gives enough accuracy for most practice,
- is specially convenient for balanced one-way data, where v/k is an integer, usually small,
- is apt for interpolation for most other cases (cases with $v < k$ deserve special - - and skeptical - - consideration anyway).

---

exhibit 3 to 5

about here

For small values of k, cases with n an integer multiple cf $k-1$ (and not k) are rather frequent (e.g. simple two-way tables), accordingly exhibit 6 offers similar tabulations for this case.

June 11, 1990

exhibit 3

Coefficients (to multiply individual standard errors) for severe simultaneous-5% SCL's and SPL's

PART 1: $1 \le k \le 9$, and $v/k = 1$, $(1.2, 1.5, 2)$, $3, 4, 6, 8, 12, 24, \infty$; upper values for SCL's, lower values for SPL's; interpolate for v/k linearly in $24k/v$

k	v/k=1	v/k=1.2	v/k=1.5	v/k=2	v/k=3	v/k=4	v/k=6	v/k=8	v/k=12	v/k=24	v/k=∞
1	±12.71			±4.30		±3.18	±2.77	±2.45	±2.31	±2.18	±2.06
2	±6.21	±5.19	±4.18	±3.50	±2.97	±2.75	±2.56	±2.47	±2.39	±2.31	±2.24
	±3.04		±2.25	±1.96	±1.73	±1.63	±1.54	±1.50	±1.46	±1.42	±1.39
3	±4.86	±4.26	(±3.72)	±3.29	±2.93	±2.78	±2.64	±2.57	±2.51	±2.45	±2.39
	±3.43			±2.32	±2.07	±1.99	±1.89	±1.83	±1.78	±1.73	±1.69
4	±4.31	±3.89	±3.52	±3.21	±2.93	±2.81	±2.70	±2.65	±2.60	±2.55	±2.50
	±3.43		±2.73	±2.42	±2.23	±2.13	±2.03	±1.99	±1.95	±1.91	±1.87
5	±4.03	±3.71		±3.17	±2.95	±2.85	±2.75	±2.70	±2.66	±2.62	±2.57
	±3.38			±2.53	±2.32	±2.23	±2.14	±2.10	±2.06	±2.02	±1.98
6	±3.86	±3.60	±3.36	±3.15	±2.96	±2.88	±2.79	±2.75	±2.71	±2.68	±2.64
	±3.32		±2.80	±2.58	±2.39	±2.30	±2.22	±2.18	±2.15	±2.11	±2.08
7	±3.75	±3.63		±3.15	±2.98	±2.90	±2.83	±2.79	±2.76	±2.72	±2.69
	±3.28			±2.62	±2.44	±2.36	±2.29	±2.26	±2.22	±2.18	±2.15
8	±3.68	±3.48	±3.31	±3.15	±3.00	±2.93	±2.86	±2.83	±2.80	±2.76	±2.73
	±3.24		±2.83	±2.64	±2.49	±2.41	±2.34	±2.31	±2.27	±2.24	±2.21
9	±3.62	±3.45		±3.15	±3.01	±2.95	±2.90	±2.86	±2.83	±2.80	±2.77
	±3.22			±2.67	±2.52	±2.45	±2.38	±2.35	±2.32	±2.29	±2.26
(24k/v =)	(24)	(20)	(16)	(12)	(8)	(6)	(4)	(3)	(2)	(1)	(0)
[6k/v]	[6]	[5]	[4]	[3]		[2]	[1]				[0]

NOTES Values are based on Bonferroni - - weakening of exception counting to "presence" vs "absence" rates instead of "per— " error rates would shorten intervals somewhat

Upper values are one-sided $(2.5/k)$% points of Student's t on v degrees of freedom

Lower values are one-sided $(5/k(k-1))$% points of $t_v\sqrt{2}$.

SP coefficients are for equal variance; a 2% increase in them covers variance ratios up to about 3 to 1; for larger ratios or guaranteed conservatism, see exhibit 14.

exhibit 4

Coefficients (to multiply individual standard errors)
for severe simultaneous-5% SCL's and SPL's

PART 2: k = 10, 20, 40, 80, 100, and v/k = 1, 1.5, 2, 3, 4, 6, 12, and ∞; upper values
for SCLs, lower values for ICLs; interpolate for k linearly in log k, for v/k
linearly in $24k/v$, or in $12k/v$, or in $6k/v$.

k	log k	$v/k=1$	$v/k=1.2$	$v/k=1.5$	$v/k=2$	$v/k=3$	$v/k=6$	$v/k=\infty$
10	1.00	±3.58	±3.43	±3.29	±3.15	±3.03	±2.91	±2.81
		±3.19	±3.01	±2.84	±2.69	±2.55	±2.42	±2.31
20	1.30	±3.45	±3.38	±3.30	±3.23	±3.16	±3.09	±3.02
		±3.13	±3.02	±2.92	±2.83	±2.74	±2.66	±2.58
40	1.60	±3.47	±3.43	±3.39	±3.35	±3.31	±3.27	±3.22
		±3.16	±3.10	±3.04	±2.98	±2.93	±2.88	±2.83
80	1.90	±3.56	±3.54	±3.51	±3.49	±3.47	±3.44	±3.42
		±3.25	±3.22	±3.18	±3.15	±3.12	±3.08	±3.05
100	2.00	±3.60	±3.58	±3.56	±3.54	±3.52	±3.50	±3.48
		±3.29	±3.26	±3.23	±3.21	±3.18	±3.15	±3.12
$(6k/v)$		[6]	[5]	[4]	[3]	[2]	[1]	[0]

NOTES: Values are based on Bonferroni - - weakening of exception counting to "presence" vs
"absence" would shorten intervals somewhat

Upper values are one-sided $(2.5/k)\%$ of Student's t on v degrees of freedom

Lower values are one-sided $(5/k(k-1))\%$ points of $t_v/\sqrt{2}$.

SP coefficients are for equal variance; a 2% increase in them covers variance ratios up
to about 3 to 1; for larger ratios or guaranteed conservatism, see exhibit 14.

exhibit 5

Coefficients (to multiply individual standard errors)
for severe simultaneous-5% SCL's and SPL's

PART 3: $k = 100, 200, 400, 800, 1000, 10^4, 10^5, 10^6 (10^9, 10^{12})$ and $v/k = 1$, and ∞; upper
values for SCL's; lower values for SPL's; interpolate for v/k linearly in k/v.

k	$v/k=1$	$v/k=\infty$
100	±3.60	±3.48
	±3.29	±3.12
200	±3.73	±3.66
	±3.43	±3.33
400	±3.87	±3.84
	±3.58	±3.52
800	±4.02	±4.00
	±3.74	±3.71
1000	±4.07	±4.06
	±3.79	±3.77
10^4	±4.57	±4.56
	±4.32	±4.32
10^5	±5.03	±5.03
	±4.81	±4.81
10^6	±5.45	±5.45
	±5.26	±5.26
(10^9)	(±6.57)	(±6.57)
	(±6.43)	(±6.43)
(10^{12})	(±7.53)	(±7.53)
	(±7.41)	(±7.41)
$[k/v =]$	[1]	[0]

Notes: Values are based on Bonferroni - - weakening of exception counting to "presence" vs
"absence" would shorten intervals somewhat.

Upper values are one sided $(2.5/k)\%$ points of Student's t on v degrees of freedom

Lower values are one-sided $(5/k(k-1))\%$ points of $t_v/\sqrt{2}$.

SP coefficients are for equal variance; a 2% increase in them covers variance ratios up to
about 3 to 1; for larger ratios or guaranteed conservatism, see exhibit 14.

---
exhibit 6

about here
---

7. Relation of SPL to SCL

There are two or three natural ways to compare SCL's with the corresponding SPL's. These include the ratio of interval lengths, SPL/SCL, and the P-value obtained when the corresponding SCL is used as if it were an SPL. The numbers are sketched in exhibits 7 and 8, respectively.

---
exhibits 7 and 8

about here

The way we look makes a considerable difference. Table 7 shows ratios mainly above .8, and increasing as we move to large k. Table 8 shows P-values mostly less than 1% (where the nominal was 5%) and decreasing as either k or v/k grows, with 0.5% reached for v/k large. This latter comparison is clearly more meaningful. We need to be careful about the distinction between SCL's and SPL's.

8. Tables for other %

If we replace 95% by Q%, hence 5% by (100-Q)%, the formulas become

$$Q\% \text{ severe SCL} \quad \text{at} \quad \pm t_v \, [(100-Q)/2k\,\%]*s$$
$$Q\% \text{ severe SPL} \quad \text{at} \quad \pm t_v \, [(100-Q)/k\,(k-1)] *s \,/\sqrt{2}$$

Notice here that these results are well defined when Q is zero or even negative. Tables 9, 10 and 11 outline the behavior for

Q% = 50%   (when we expect an average of one exception for every two batches or families)
Q% = 0%   (when we expect an average of one for one)
Q% = -100%   (when we expect an average of two for one)

The one numerically interesting fact about these tables is that SPL's do, in a few instances, get somewhat larger than the corresponding SCL's. We are so far from the tails that the argument toward the close of the introduction fails.

---
exhibits 9 to 11

about here

June 11, 1990

exhibit 6

Critical values (to multiply individual standard errors)
for severe simultaneous-5% SCL's and SPL's

Special case: $k = 1, 2, \ldots, 10$, $v/(k-1) = 1, 2, 3, 4, 6, 8, 12, \infty$; upper values
are for SCL's; lower values for SPL's; do not interpolate here, unless
v is a convenient multiple of $k-1$ interpolate instead in exhibit 3
(in exhibit 4 for $k \geq 10^2$)

k	$v/k^*=1$	$v/k^*=2$	$v/k^*=3$	$v/k^*=4$	$v/k^*=6$	$v/k^*=8$	$v/k^*=12$	$v/k^*=24$	$v/k^*=\infty$	$k^*=k-1$
2	±12.71	±4.30	±3.18	±2.77	±2.45	±2.31	±2.18	±2.06	±1.96	1
	±8.98	±3.04	±2.25	±1.96	±1.73	±1.63	±1.54	±1.45	±1.39	
3	±6.21	±3.50	±2.97	±2.75	±2.56	±2.47	±2.39	±2.31	±2.24	2
	±5.41	±2.80	±2.32	±2.13	±1.97	±1.89	±1.82	±1.75	±1.69	
4	±4.86	±3.29	±2.93	±2.78	±2.64	±2.57	±2.51	±2.45	±2.39	3
	±4.41	±2.73	±2.38	±2.23	±2.09	±2.03	±1.97	±1.92	±1.87	
5	±4.31	±3.21	±2.93	±2.81	±2.70	±2.65	±2.60	±2.55	±2.50	4
	±3.96	±2.71	±2.42	±2.30	±2.19	±2.13	±2.08	±2.03	±1.98	
6	±4.03	±3.17	±2.95	±2.85	±2.75	±2.70	±2.66	±2.62	±2.57	5
	±3.71	±2.70	±2.46	±2.36	±2.25	±2.21	±2.16	±2.12	±2.08	
7	±3.86	±3.15	±2.96	±2.88	±2.79	±2.75	±2.71	±2.68	±2.64	6
	±3.56	±2.71	±2.50	±2.40	±2.31	±2.27	±2.23	±2.19	±2.15	
8	±3.75	±3.15	±2.98	±2.90	±2.83	±2.79	±2.76	±2.72	±2.69	7
	±3.45	±2.72	±2.53	±2.44	±2.36	±2.32	±2.28	±2.24	±2.21	
9	±3.68	±3.15	±3.00	±2.93	±2.86	±2.83	±2.80	±2.76	±2.73	8
	±3.38	±2.73	±2.56	±2.48	±2.40	±2.36	±2.33	±2.29	±2.26	
10	±3.62	±3.15	±3.01	±2.95	±2.89	±2.86	±2.83	±2.80	±2.77	9
	±3.33	±2.74	±2.58	±2.51	±2.44	±2.40	±2.37	±2.34	±2.31	
$24k^*/v$	24	12	8	6	4	3	2	1	0	

NOTES: Values are based on Bonferroni - - weakening of exception counting to "presence" vs "absence"
would shorten intervals somewhat.

Upper values are one sided $(2.5/k)\%$ of Student's t on v degrees of freedom

Lower values are one-sided $(5/k(k-1))\%$ points of $t_v/\sqrt{2}$.

SP coefficients are for equal variances; a 2% increase in them covers variance ratios up to
about 3 to 1; for larger ratios or guaranteed conservatism, see exhibit 14.

exhibit 7

Ratio of (severe) 95% SPL to (severe) 95% SCL

k	$v/k=1$	$v/k=2$	$v/k=\infty$
2	.49	.56	.62
3	.71	.71	.71
4	.80	.77	.75
5	.84	.80	.77
10	.89	.85	.82
20	.91	.88	.85
40	.91	.89	.88
100	.91	.91	.90
1000	.93	.93	.93
10000	.95	.95	.95

NOTE: Ratio of IPL to ICL is always 0.707.

Ratios above are for the equal-variance case.

exhibit 8

Percentage levels (tail areas) obtained by using 5% severe SCL's as severe SPL's

k	$v/k = 1$	$v/k = 2$	$v/k = 3$	$v/k = 6$	$v/k = \infty$
3	1.89%	1.05%	.75%	.46%	.22%
4	2.20%	1.14%	.82%	.50%	.24%
6	2.36%	1.18%	.83%	.53%	.28%
10	2.20%	1.09%	.78%	.54%	.32%
20	1.78%	.88%	.67%	.50%	.37%
40	1.24%	.72%	.59%	.48%	.41%
100	.83%	.60%	.54%	.48%	.43%
200	.68%	.56%	.53%	.49%	.45%
400	.62%	.53%	.50%	.47%	.45%
1000	.57%	.52%	.51%	.49%	.47%

NOTE: Tail areas above are for the equal-variance case.

exhibit 9

Coefficients for severe 50% SPL's (lower entries) and SCL's (upper entries)

k	$\log k$	$v/k = 1$	$v/k = 1.5$	$v/k = 3$	$v/k = \infty$
2	.30	1.60	1.42	1.27	1.15
		0.58	0.54	.51	0.48
3	.48	1.82	(1.66)	1.50	1.38
		1.29	(1.17)	1.06	0.98
4	.60	1.94	1.78	1.65	1.53
		1.62	1.47	1.34	1.22
5	.70	2.01	(1.88)	1.75	1.64
		1.82	(1.65)	1.51	1.39
6	.84	2.07	1.95	1.83	1.73
		1.94	1.77	1.63	1.50
8	.90	2.16	2.05	1.95	1.86
		2.10	1.94	1.80	1.67
10	1.00	2.23	2.13	2.04	1.96
		2.20	2.04	1.91	1.80
20	1.30	2.42	2.36	2.30	2.24
		2.43	2.32	2.22	2.13
40	1.60	2.62	2.58	2.54	2.50
		2.62	2.55	2.48	2.41
80	1.90	2.81	2.78	2.76	2.73
		2.80	2.76	2.71	2.67
125	2.10	2.93	2.91	2.90	2.88
		2.92	2.89	2.86	2.83
200	2.30	3.06	3.05	3.04	3.02
		3.05	3.03	3.00	2.98
400	2.60	3.25			3.23
		3.24			3.19
800	2.90	3.43			3.42
		3.42			3.40
1250	3.10	3.55			3.54
		3.54			3.52
2000	3.30	3.67			3.66
		3.66			3.65

NOTE: Upper values are one-sided $t_v[25\%/k]$

Lower values are $t_v[50\%/k(k-1)]/\sqrt{2}$

SP coefficients are for equal variance; a 2% increase in them covers variance ratios up to about 3 to 1; for larger ratios or guaranteed consideration, see exhibit 14.

exhibit 10

Coefficients for severe "0%" SCL's (upper entries) and SPL's (lower entires)
(average of one exception per batch or family)

k	$v/k =1$	$v/k =1.5$	$v/k =3$	$v/k =\infty$
2	.82	.76	.72	.67
	.00	.00	.00	.00
3	1.15	–	–	.97
	.81	–	–	.68
4	1.34	–	–	1.15
	1.19	–	–	.98
5	1.48	(1.41	1.34	1.28
	1.43	(1.33)	1.24	1.16
6	1.57	–	–	2.29
	1.46	–	–	1.30
8	1.71	–	–	1.53
	1.78	–	–	1.48
10	1.81	1.75	1.70	1.64
	1.91	1.80	1.71	1.62
20	2.09	2.04	2.00	1.96
	2.21	2.13	2.05	1.97
40	2.32	–	–	2.29
	2.45	–	–	2.28
80	2.56	–	–	2.50
	2.66	–	–	2.55
125	2.70	–	–	2.65
	2.79	–	–	2.70
200	2.84	2.83	2.82	2.81
	2.93	2.91	2.89	2.97
$[3k/v]$	[3]	[2]	[1]	[0]
$[k/v]$	[1]			[0]

Upper lines : SCL = one-sided $50\%/k$ point of t_v

Lower lines : SPL = one-sided $100\%/k(k-1)$ point of t_v divided by $\sqrt{2}$

SP coefficients are for equal variance; a 2% increase in them covers
variance ratios up to about 3 to 1; for larger ratios or guaranteed conservatism,
see exhibit 14.

exhibit 11

Coefficients for severe "minus 100%" SCL's (upper entries) and SPL's (lower entries)
(average of two exceptions per batch or family)

k	$v/k = 1$	$v/k = 2$	$v/k = \infty$
2	.00	.00	.99
	–	–	–
3	.47	.45	.43
	.34	.32	.30
4	.74	–	–.67
	.78	–	–
5	.92	.88	.84
	1.04	.97	.91
6	–	–	.97
	–	–	1.06
8	–	–	1.15
	–	–	1.27
10	1.37	1.13	1.28
	1.62	1.52	1.42
20	1.72	1.68	1.64
	2.00	1.90	1.81
40	2.02	1.99	1.96
	2.28	2.20	2.13
80	2.28	–	–
	2.52	–	–
125	2.44	–	–
	2.66	–	–
200	2.60	2.59	2.58
	2.81	2.78	2.75

Upper values are one-sided $[100\%/k]$ points of t_v

Lower values are $[200\%/k\,(k-1)]$ points of $t\sqrt{\sqrt{2}}$

SP coefficients are for equal variance; a 2% increase in them covers variance ratios
up to about 3 to 1; for larger ratios or guaranteed conservatism, see exhibit 14.

9. Tight simultaneous limits

As remarked generally above, tight 5% SPL's, which guarantee any exceptional comparisons in at most an average of once in 20 applications, are calculated as

$$\text{(tight) 5\% SPL} = \frac{1}{2} q_v(k/5\%) * s_v$$

where q_v is the distribution of the Studentized range of k observations where the Studentization is by s_v, which is on v (Gaussian) degrees of freedom.

We have already given tables for severe 5% SPL's. Thus it seems reasonable to now tabulate the difference

coefficient of s_v in a *severe* 5% SPL — coefficient of s_v in a *tight* 5% SPL

as in exhibit 12.

---

exhibit 12

<u>about here</u>

The values of these differences are not all very small, especially for v/k small. On the other hand, for $v/k \geq 3$; or $k \geq 1.5$; and $v/k = 2$, we have

$$.07 \leq \text{difference} \leq .17$$

so that, in this range of v/k, diversity of values is relatively small - - with a difference of .12 as a simple rough approximation.

If we want a *closer* approximation, $.08 + .084k/v^{3/2}$ is good to ± 0.03 except for 8 entries with both k and v/k small, as indicated on the exhibit.

The corresponding relations for SCL's involve

$$\text{("tight") 5\% SCL} = |m|_{k,v}^{.05}$$

= the upper 5% of the Studentized maximum modulus for k candidates and v degrees of freedom.

Again, it seems helpful to tabulate the corresponding difference, as in exhibit 13.

---

exhibit 13

<u>about here</u>

June 11, 1990

exhibit 12

Coefficient of severe 5% SPL MINUS coefficient of tight 5% SPL

A) Differences of coefficients

k	$v/k = 1$	$v/k = 1.2$	$v/k = 1.5$	$v/k = 2$	$v/k = 3$	$v/k = 6$	$v/k = \infty$
2	.00	.00	.00	.00	.00	.00	.00
3	.47	(.35)	(.23)	.15	.10	.09	.03
4	.55	(.40)	.28	.20	.13	.08	.05
5	.54	.40	(.30)	.20	.14	.09	.05
6	.51	(.35)	.29	.20	.14	.09	.06
8	.44	(.35)	.27	.19	.15	.10	.07
10	.39	.31	.24	.19	.14	.10	.07
15	.32	.26	.21	.17	.13	.10	.07
20	.27	.22	.18	.15	.12	.10	.07
30	.22	.20	.16	.14	.12	.09	
40	.19	.17	.14	.12	.11	.09	.08
70	.14	.13	.12	.11	.11	.09	.08
100	.13	.12	.12	.11	.10	.09	.08
$[6k/v]$	[6]	[5]	[4]	[3]	[2]	[1]	[0]

NOTES: Interpolate linearly in $6k/v$, and linearly in $\log k$
Errors of $\pm$ .01 or $\pm$ .02 are not unlikely.

Values given are for equal variances; a 2% increase in the resulting narrow
SP coefficients covers variance ratios up to about 3 to 1; for larger ratios or
guaranteed conservatism, see exhibit 14.

Entries well approximated by $.08 + .057(6k/v)^{3/2}/k = .08 + .084k/v^{3/2}$

B) Above differences MINUS $.08 + .084k/v^{3/2}$ (values not given are .00 or $\pm$ .01)

k	$v/k = 1$	$v/k = 1.2$	$v/k = 1.5$	$v/k = 2$	$v/k = 3$	$v/k = 6$	$v/k = \infty$
3	-.12			-.10	-.07	-.02	-.05
4	.03		-.03	-.03	-.03	-.03	-.03
5	.06	.04				-.02	-.03
6	.09		.02				.-.02
8	.06		.03				
10	.05	.03	.02				
15	.02	.02		.02			
20		.02					
30							
40	-.02						
70	-.04	-.03					
100	-.03	-.02					

exhibit 13

Coefficient of severe 5% SCL MINUS coefficient of tight 5% SCL

A) Differences themselves

k	$v/k = 1$	$v/k = 1.2$	$v/k = 1.5$	$v/k = 2$	$v/k = 3$	$v/k = 6$	$v/k = \infty$
2	.00	.00	.00	.00	.00	.00	.00
3	.70	(.34)	.20	.08	.04	.02	.00
4	.37	.22	.12	.07	.04	.02	.01
5	.25	.15	.10	.07	.04	.02	.01
6	.18	.14	.09	.06	.03	.02	.01
7	.15	.12	.08	.08	.03	.02	.01
8	.14	.10	.07	.05	.03	.02	.01
9	.13	.10	.07	.05	.03	.02	.01
10	.11	.08	.06	.03	.02	.01	.01
15	.05	.05	.04	.03	.02	.01	.01
20	.05	.04	.04	.03	.02	.01	.01
> 20	$1/k$	$.8/k$	$.6/k$	$.4/k$	$.2/k$	$.1/k$	.01
$[6k/v]$	[6]	[5]	[4]	[3]	[2]	[1]	[0]

NOTES   Interpolate in $6k/v$
Errors of $\pm .01$ or $\pm .02$ are not unlikely.

Well approximated by $.095(6k/v)^{4/3}/k = .109k^{1/3}/v^{4/3} = .109(k/v)^{1/3}$

B) Above differences MINUS $.109k^{1/3}/v^{4/3}$

k	$v/k = 1$	$v/k = 1.2$	$v/k = 1.5$	$v/k = 2$	$v/k = 3$	$v/k = 6$	$v/k = \infty$
3	.35	.02	.00	-.02	-.04		
4	.11	-.01	-.03	-.01	-.02		
5	.04	-.01	-.02	.00	-.01		

Remaining differences MINUS $.109k^{1/3}/v^{4/3}$ are .00 or $\pm .01$

Now we see that, so long as $v/k \geq 2$ ($v/k \geq 1.5$ for $k \geq 5$) we will have

$$.00 \leq \text{difference} \leq .10$$

Thus the difference between "severe" and "tight" is typically even smaller for SCL's than for SPL's.

If we want a closer proximation, $.109k^{1/3}/v^{4/3}$ is good to ± 0.01 except for small k combined with small v/k.

10. Equal or unequal variances

Throughout, SP coefficients have been defined and tabulated for the case of equal variance. If variances associated with the k candidates differ substantially, small increases in the SP coefficients are needed. This is because the standard error of a difference is proportional to the square root of the sum of squares of the standard errors concerned rather than to their sum. Exhibit 14 shows the ratios of increase needed (a) for conservatism, and (b) for roughly maintaining the simultaneous exception rates in an average situation.

---
exhibit 14

about here

If one difference, or a few differences, are of special importance, we may use the exact factor

$$\frac{\sqrt{2(1 + \text{ratio})}}{1 + \sqrt{\text{ratio}}}$$

where "ratio" is the ratio of variances, to correct the sum of the two raw SPL's.

11. Severe or tight

Severe ($=$ Bonferroni) limits are considered here as precise answers to a different, more challenging, question, rather than as approximations to, or outer bounds for, the tight ($=$ Studentized range) limits. (If a rough approximation to the tight limits is required, subtracting 0.12 from the SCL coefficient or 0.05 from the SPL coefficient will of en be good enough.) The severe limits, because they involve only the distribution of Student's t, which is now widely available on computers and sophisticated hand-held calculators alike, are available for any combination of parameters we may desire. The tight SPL limits, using the Studentized range distribution are not as available, though tables let us cover $k \leq 100$ as needed. The tight SCL limits, using the Studentized maximum modulus distribution, are still less available, with only limited tables (Nemenyi 1953, Pillai and Ramachandram 1954) visible.

June 11, 1990

exhibit 14

Adjustment factors for SP coefficients
to allow for unequal variance

Ratio of variances	Factor for conservatism*	Factor to preserve exception rates**
2 : 1	1.015	1.008
3 : 1	1.035	1.018
4 : 1	1.054	1.024
5 : 1	1.071	1.036
6 : 1	1.085	1.042
7 : 1	1.094	1.047
8 : 1	1.108	1.054
9 : 1	1.118	1.059
10 : 1	1.127	1.064

*Required for $k = 2$

**Convenient approximation, based on a number of spot checks

If a few variances are unusually large, a simple, conservative approach is to use SCL's as SPL's for those candidates and select a factor for the other SPL's based on the extreme ratio of the remaining variances.

Standardization on severe limits thus seems a desirable choice, with a caveat that we can shorten the intervals slightly, when desired, by asking less of our intervals and switching to the tight limits, with the aid of either exhibits 12 and 13 or the "close approximation" formulas.

REFERENCES

Harter, H. S. (1961). "Use of tables of percentage points of range and Studentized range," *Technometrics*, **3**: 407-411.

Harter, H. S. (1969). *Order Statistics and Their Use in Testing and Estimation, Volume 1. Tests based on Range and Studentized Range of Samples from a Normal Population*, Aerospace Research Laboratories, USAF. 761pp.

Miller, R. G. (1981). *Simultaneous Statistical Inference*, New York: McGraw Hill (1966), (2nd ed., New York: Springer-Verlag).

Nemenyi, P. (1953). "Percentage points of the Studentized maximum modulus," *Memorandum Report 48, Statistical Research Group*, Princeton University, Princeton, NJ.

Pillai, V. C. S., and Ramachandram, K. V. (1954). "On the distribution of the ratio of the *i*th observation in an ordered sample from a normal population to an independent estimate of the standard deviation, *Ann. Math. Statist.* **25**: 565-572.